EPH165

Race, social determinants of health (SDoH), and stage at breast cancer diagnosis in electronic health records (EHR) data

Ellen Stein, MS[1], Zuzanna Drebert, PhD[1], Katelyn J White, MPH[1], Julia A O'Rourke, PhD[1], Marley Boyd, MS[1], E. Susan Amirian, PhD[1]
[1]TriNetX, LLC., Cambridge, MA, USA

BACKGROUND

- Early diagnosis of breast cancer is associated with improved outcomes.
- Black women are generally more likely than White women to be diagnosed at a later stage.
- Some studies have found that racial disparities in diagnostic timing persist despite adjustment for SDoH and other factors.

OBJECTIVE

We aim to explore the associations between race, SDoH, and stage at breast cancer diagnosis using EHR data.

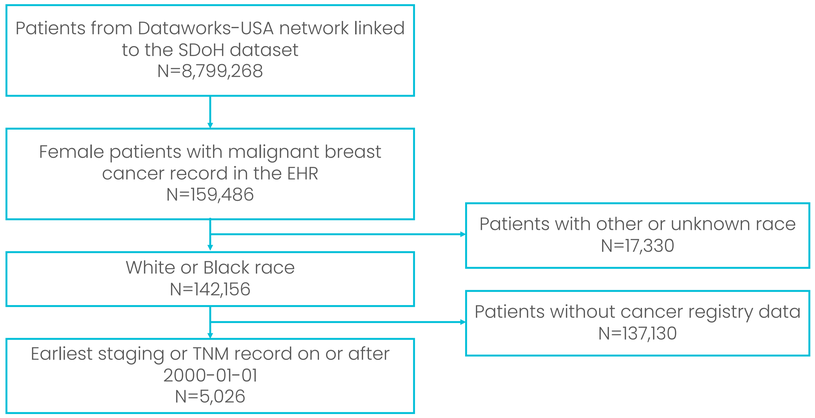

Figure 1. Patient cohort selection

METHODS

- In this study we utilized deidentified EHR data from 9 healthcare organizations (HCOs) belonging to the TriNetX Dataworks-USA federated research network.
- The de-identified EHR data were linked through tokenization to third-party SDoH information to explore the relationships between race, SDoH, and stage at diagnosis (early stage: 0-2; late: 3-4).
- Female Black and White patients diagnosed with breast cancer between 2000-2024 were included.
- Descriptive statistics (e.g., standardized mean differences, SMD) were examined.
- We used logistic regression to examine associations between race and late-stage diagnosis, with White race serving as the referent group.

Table 1. Baseline Characteristics

CHARACTERISTIC		OVERALL N=5026	BLACK N=936	WHITE N=4090	SMD
Cancer stage at diagnosis	Early stage (0-2)	4449 (88.5)	797 (85.1)	3652 (89.3)	0.124
	Late stage (3-4)	577 (11.5)	139 (14.9)	438 (10.7)	
Age at diagnosis					
mean (SD)		58.4 (12.0)	57.3 (12.3)	58.7 (11.9)	0.115
median [Q1,Q3]		59.0 [50.0,67.0]	57.0 [49.0,66.0]	59.0 [50.0,67.0]	0.115
[min, max]		[19.0,90.0]	[21.0,90.0]	[19.0,90.0]	0.115
Ethnicity[2], n (%)	Hispanic or Latino	30-40 (0.7)	<10 (0.5)	30-40 (0.9)	0.233
	Not Hispanic or Latino	4620-4230 (91.9)	810-820 (87.1)	3800-3810 (93.0)	
	Unknown	370-380 (7.4)	110-120(12.4)	250-260 (6.1)	
Language[2], n (%)	English	3500-3510 (69.7)	620-630 (66.3)	2870-2890 (70.4)	0.129
	Non-English	30-40 (0.7)	<10 (0.5)	30-40 (0.9)	
	Unknown	1480-1490 (29.6)	310-320 (33.2)	1170-1180 (28.7)	
Education, n (%)	Attended Vocational/Technical	64 (1.3)	10 (1.1)	54 (1.3)	0.32
	Completed College	1288 (25.6)	192 (20.5)	1096 (26.8)	
	Completed Graduate School	1344 (26.7)	184 (19.7)	1160 (28.4)	
	Completed High School	1140 (22.7)	250 (26.7)	890 (21.8)	
	Unknown (missing information)	621 (12.4)	165 (17.6)	456 (11.1)	
	Some College	569 (11.3)	135 (14.4)	434 (10.6)	
Region (US)[2], n (%)	Midwest	1420-1430 (28.3)	130-140 (14.5)	1290-1300 (31.5)	0.433
	Northeast	80-90 (1.7)	<10 (0.5)	70-80 (1.8)	
	South	3380-3390 (67.5)	760-770 (81.5)	2620-2630 (64.3)	
	West	130-140 (2.6)	30-40 (3.5)	90-100 (2.4)	
Estimated household income, n (%)	Less than $20,000	434 (8.6)	201 (21.5)	233 (5.7)	0.614
	$20,000-$49,999	751 (14.9)	209 (22.3)	542 (13.3)	
	$50,000-$99,999	1473 (29.3)	246 (26.3)	1227 (30.0)	
	$100,000+	2368 (47.1)	280 (29.9)	2088 (51.1)	
BMI, mean (SD)		29.5 (6.6)	32.4 (6.9)	29.0 (6.4)	-0.509
Obesity record, n (%)	0	1316 (26.2)	130 (13.9)	1186 (29.0)	0.375
	1	940 (18.7)	203 (21.7)	737 (18.0)	
	Unknown	2770 (55.1)	603 (64.4)	2167 (53.0)	
Smoker record, n (%)	1	255 (5.1)	45 (4.8)	210 (5.1)	0.015
	Unknown	4771 (94.9)	891 (95.2)	3880 (94.9)	
Rural-urban commuting area codes (RUCA), n (%)	Urban	3545 (70.5)	730 (78.0)	2815 (68.8)	0.223
	Suburban	672 (13.4)	84 (9.0)	588 (14.4)	
	Small town	265 (5.3)	35 (3.7)	230 (5.6)	
	Rural	454 (9.0)	69 (7.4)	385 (9.4)	
	Not coded or Unknown	90 (1.8)	18 (1.9)	72 (1.8)	

[2]Patient counts are shown as ranges when at least one subcategory includes fewer than 10 patients.

Table 2. Univariable and Multivariable associations between race and stage at breast cancer diagnosis.

CHARACTERISTIC	LATE-STAGE DIAGNOSIS	
	Odds Ratio (95% CI)	Adjusted Odds Ratio (95% CI)[3]
Race White Black	Ref 1.45 (1.18, 1.79)	Ref 1.32 (1.05, 1.64)

[3]Adjusted for age at diagnosis, income, obesity, education, and RUCA

RESULTS

Differences in SDoH were observed between Black and White patients with breast cancer.

- Relative to White women, a higher proportion of Black women were in the lower household income category (<$50,000 annually; SMD 0.614).
- Black women had higher mean BMI at diagnosis than White women (32.4 vs 29.0, respectively; SMD -0.509).
- There were differences in highest level of educational attainment (SMD 0.32).

There were also differences observed between Black and White patients who had a late stage a diagnosis.

- Among those with a lower household income, a higher proportion of Black women had a late stage at diagnosis (18.5%) than White women (12.0%).
- A higher proportion of Black women had late-stage diagnosis among both educational attainment strata.

In unadjusted logistic regression analysis, Black race was significantly associated with increased odds of late-stage breast cancer diagnosis compared to White race (odds ratio 1.45, 95% CI: 1.18–1.79).

After adjusting for age, income, obesity, and rural-urban commuting area codes (RUCA) the association between race and stage at diagnosis was attenuated but remained statistically significant (odds ratio: 1.32, 95% CI 1.05-1.64).

CONCLUSION

- EHR data linked with supplementary data can be leveraged to elucidate racial disparities in timely cancer diagnosis.
- Black patients had higher odds of late-stage breast cancer diagnosis than White patients, even after adjusting for age, income, obesity, and RUCA.

Download Poster PDF

ISPOR 2025 | Montréal, QC, Canada
Design services were provided by Kelly Marsocci and were funded by TriNetX, LLC.